A Modulo Packet Marking Approach to Protect Cloud Environment against DDoS Attacks

E.Anitha and Dr.S.Malliga

Abstract— Cloud computing uses internet and remote servers for maintaining data and applications. It offers through internet the dynamic virtualized resources, bandwidth and on-demand software's to consumers and promises the distribution of many economical benefits among its adapters. It helps the consumers to reduce the usage of hardware, software license and system maintenance. Simple Object Access Protocol (SOAP) is the system that allows the communications interaction between different web services. SOAP messages are constructed using either HyperText Transport Protocol (HTTP) and/or eXtensible Mark-up Language (XML). Cloud computing suffers from major security threat problem by Denial of Service (DoS) attacks which are more difficult to defend. The attackers in Distributed DoS (DDoS) attacks modify the source addresses in the attack packets to hide their identity, and making it difficult to distinguish such packets from those sent by legitimate users. This idea is called IP address spoofing. They are intentionally sent to flood and destroy the communication channel of the cloud service provider. To address the problem of DDoS attacks against cloud web services there is a need to distinguish between the legitimate and illegitimate messages. This can be done by using the rule set based detection, called CLASSIE and modulo marking method is used to avoid the spoofing attack. Reconstruct and Drop method is used to make decision and drop the packets on the victim side. It enables us to improve the reduction of false positive rate and increase the detection and filtering of DDoS attacks.

Index Terms— Cloud Computing, Cloud Security, Denial of Service, Intrusion Detection and Traceback

1 Introduction

Cloud computing is a new computing model in which resources are pooled to provide software, platform and infrastructure to as many users as possible by sharing the available resources. In this model "customers" plug into the "cloud" to access IT resources which are priced and provided "on-demand". The NIST (US National Institute of Standards and Technology) definition of cloud computing is " a model for enabling ubiquitous, convenient, on-demand network access to shared pool of configurable computing resources(e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

1.1 Hall Marks Of Cloud

On-demand self service, broadband network access, resource pooling, rapid elasticity are some of the essential characteristics of the cloud model. The cloud can be deployed for private, public, community or uses. Private cloud will be used by an organization and its customers, whereas public cloud is made available for public use. Community model is for a community of users having same mission/goal. Hybrid model of cloud shares the properties of any of the above models.

Shabeeb et al (2012) discussed about the cloud services. The cloud delivers its services in the form of software, platform and infrastructure. Costly applications like ERP, CRM will be offloaded onto the cloud by provider. They run at providers cost. Platform includes the languages, libraries etc. and the database, operating system, network bandwidth comes under infrastructure.

1.2 Security Issues

Trustworthiness of the cloud service provider is the key concern. The organizations are deliberately offloading their sensitive as well as insensitive data to cloud for getting the services. The cloud works on pay for use basis. If numerous requests are sent to a server on cloud by the DoS attacker, the owner of that particular cloud have more requests for process. Moreover, other users will be denied of the service which they request as the server on cloud is expending all its requests for serving the malicious DoS request. The situation will be more drastic if the attacker compromises some more hosts for sending the flood request, which is called DDoS.

Chonka et al (2011) discussed the variant forms of DDoS attack tools like Agobot, Mstream and Trinoo which are still used by attacker today. But, most attackers are more inclined to use the less complicated web based attack tools like Extensible XML-based Denial of Service (X-DoS) and HTTP-based Denial of Service (H-DoS) attack due to their simple implementation and lack of any real defences against them.

The rest of the report is organized as follows: In chapter 2, related works are reviewed. Chapter 3 explains the proposed method to overcome the problem of existing method. Chapter 4 focuses on system implementation. Chapter 5 provides the conclusion and future work.

2 Literature Review

A DoS attack is designed to prevent legitimate access to a resource. In the context of the Internet, an attacker can "flood" a victim's connection with random packets to prevent legitimate packets from getting through. These internet Denial of Service attacks have become more prevalent recently due to their near untraceability and relative ease of execution.

Dos attacks are so difficult to trace because the only hint a victim has, is the source of a given packet which can be easily forged. Dean et al (2001) presented a solution to the problem of determining the path a packet traversed over the Internet (called the traceback problem). It reframes the traceback problem as a polynomial reconstruction and uses algebraic techniques from coding theory and learning theory to provide robust methods of transmission and reconstruction.

Savage et al (2001) presented an approach to the traceback problem that addresses the needs of both victims and network operators. The possibility of tracing flooding attacks by "marking" packets, either probabilistically or deterministically, with the addresses of the routers they traverse. The victim uses the information in the marked packets to trace an attack back to its source. A router "marks" one or more packets by augmenting them with additional information about the path they are travelling. The victim attempts to reconstruct the attack path using only the information in the marked packets. It allows a victim to identify the network path(s) traversed by attack traffic without requiring interactive operational support from Internet Service Providers (ISPs).

Belenky et al (2003) proposed a Deterministic Packet Marking (DPM), a new approach to IP traceback. The 16-bit Packet ID Field and 1-bit Reserved Flag (RF) in the IP header will be used to mark packets. The packet is marked by the interface closest to the source of the packet. A general principle in handling DDoS attacks is to rely only on the information transferred in the DPM mark. The DPM Mark can be used to not only transfer the bits of the ingress address but also some other information. This additional information should enable the destination to determine which ingress address segments belong to which ingress address. At the victim, a table matching the source addresses to the ingress addresses is maintained. The reconstruction procedure utilizes the data structure called Reconstruction Table (RecTbl), in which the destination would first put the address segments. After segments corresponding to the same ingress address have arrived to the destination, the ingress address for a given source address becomes available to the victim.

Xiang et al (2009) presented a Flexible Deterministic Packet Marking (FDPM) which provides a defense system with the ability to find out the real sources of attacking packets that traverse through the network. The FDPM scheme utilizes various bits (called marks) in the IP header. The mark has flexible lengths depending on the network protocols used, which is called flexible mark length strategy. The flexibility of FDPM is twofold. First, it can use flexible mark length according to the network protocols that are used in the network. This characteristic of FDPM gives it much adaptability to current heterogeneous networks. Second, FDPM can adaptively adjust its marking process to obtain a flexible marking rate. This characteristic prevents a traceback router from the overload problems. It has been used to not only trace DDoS attacking packets but also enhance filtering attacking traffic.

hoi and Dai (2004) presented a marking scheme (with marking and traceback algorithms) in which a router marks a packet with a link that the packet came through. Links of a router are represented by Huffman codes according to the traffic distribution among the links. When a router marks a packet with address information, the information is not of the router that is marking but of a router that sent the packet to the current router and it uses a special table called link table, which shows all the links between the router and its adjacent routers. The router appends to the marking field a Huffman codeword representing the link number of the link (router) through which the packet arrived.

When the marking field of a packet becomes short of space left to append the corresponding Huffman codeword for the link number, the router stores the content of the marking field with a message digest of the packet into the router's local memory, and then clears the field and appends the codeword. The stored link sequence can be retrieved via the message digest of the packet from the intermediate router during an IP traceback procedure. This scheme marks every packet, therefore IP traceback can be accomplished with only a packet unlike in probabilistic markings; also it requires far less amount of memory compared to logging methods and is robust in case of DDoS.

Chonka et al (2008c) proposed an IP traceback scheme using a machine learning technique called Intelligent Decision Prototype (IDP). IDP can be used on both Probabilistic Packet Marking (PPM) and DPM traceback schemes to identify DDoS attacks. IDP is a supervised machine learning application that is employed into two parts. The first part, called Pre-Marked Decision (PMD), is located at the edge of the routers, like DPM. If the traffic is legitimate, the packet is forwarded onto the next router or host. If PMD decides that the packet shows signs that it is not legitimate, it sends it for packet marking. The second part of IDP is made up of two sections. One section is to deal with reconstructing the path back to the source of the attack and the second section uses machine learning method, called Reconstruct And Drop (RAD), to deal with the actual attack packet. This will greatly reduce the packets that are marked and in effect make the system more efficient and effective at tracing the source of an attack compared with other methods.

Service Oriented Architecture is an architectural paradigm and discipline that may be used to build infrastructures enabling those with needs (consumers) and those with capabilities (providers) to interact via services across disparate domains of technology and ownership. Chonka et al (2008a) suggested a new approach, Service Oriented Traceback Architecture (SOTA), which provides a framework to be able to identify the source of an attack. The main objective of SOTA is to apply a SOA approach to traceback methodology, in order to identify the true source of a DDoS. SOTA is based upon a popular form of packet marking called DPM. SOTA framework employs the DPM methodology and places Service-Oriented Traceback Mark (SOTM) within a web service message.

Chonka et al (2008b) extended SOTA, in order to defend Web Services against DDoS attacks. SOTA's main objective is to identify the true identity of forged messages, since an attacker tries to hide their identity to avoid current defence systems and escape prosecution. To accomplish the main objective, SOTA should be attached as close to the source of the attack. When an incoming SOAP message comes into the router, it is tagged with own SOAP header. The header can be used to traverse the network back to the true source of the attack. Chonka et al (2009a) extended SOTA, by applying the framework to Open Grid Service Architecture (OGSA) and further introduced a defense filter called XDetector (XML Detector), in which it is distributed throughout the grid, in order to properly defend it. The XML-Based Detector is trained Back Propagation Neural Network, in order to detect and filter out Xml-Based Denial of Service (X-DoS) messages. XDetector is located before the web server in order to provide the greatest resource efficiency and protection.

Chonka et al (2011) offered a solution for DDoS attacks by the use of service oriented traceback architecture in the area of cloud computing. Cloud TraceBack (CTB) is used to find the source of the attacks, and introduced the use of a back propagation neutral network, called Cloud Protector (XDetector), which was trained to detect and filter attack traffic. In an attack scenario, the attack client will request a web service from CTB, which in turn will pass the request to the web server. The attack client will then formulate a SOAP request message based on the service description formulated by WSDL. Upon receipt of SOAP request message, SOTA will place a SOTM within the header. Once the CTBM has been placed, the SOAP message will be sent to the Web Server. Upon discovery of an attack, the victim will ask for reconstruction to extract the mark and inform them of the origin of the message. The reconstruction will also begin to filter out the attack traffic. It helps to detect and filter most of the attack messages and identify the source of the attack within a short period of time.

3 Proposed Work

As attacking another machine with a flood of messages to a point where it can only handle a few requests at a time or alternatively the system totally collapses. A new form of DDoS attacks that could potentially bring down a cloud web services are HTTP and XML DoS attacks. They are combinedly called as HX-DoS attack.

In a DDoS attack scenario, an attacker has compromised a client who has an account to access the cloud service provider server. This way they have a direct connection through the system. The attacker then installs the DoS attack program at the user end and initiates it. To distinguish between them, the first method adopts Intrusion Detection System (IDS) by using a decision tree classification system called as CLASSIE.

CLASSIE is located one hop away from host. CLASSIE's rule set has been built up over time to identify the known DDoS messages. With known DDoS attacks like XML injection or XML Payload Overload, CLASSIE is able to be trained and tested to identify these known attributes. Upon detection of DDoS message, CLASSIE drops the packet which matches the rule set. After examined by the CLASSIE, then the packets are subjected to marking. Fig.1 shows the conceptual diagram of the proposed approach.

The new marking scheme is the modulo packet marking algorithm. As the packets travel through the network, they are marked with router information using modulo technique. Upon traceback request, reverse modulo is used to reconstruct the path traversed by the packets. The marking is done on both edge and core routers. When an edge router decides to mark an incoming packet, it fetches the code to be marked that corresponds to physical address of the host from the lookup table and encodes it into the packet. The edge router requires one bit for indicating whether the packet is marked or not and few bits for marking code.

And it maintains a lookup table called MACtoID table, which has physical address of the hosts attached to the network and equivalent numeric code for each of the physical addresses.The algorithm for marking at edge router has the following steps:

Step 1: For every packet, use the physical address of the sender to find the code to be marked from MACtoID table.
Step 2: Set marked field.
Step 3: Stamp the code into marked field.
Step 4: Forward the packet to the next router.

A core router marks if only the packet has been already marked by the edge router. Otherwise it would simply forward the packets. The core router maintains a table called MACtoInterface, that contains the physical addresses of all of its hardware input interfaces and link numbers assigned to each of these interfaces. The algorithm for marking at core router is:

Step 1: For every packet, if the marked field is set, use MACtoInterface table and find the link number for the inbound interface on which the packet arrived.
Step 2: Calculate the new marking information.
Step 3: Forward the packet to the next router.

When a router decides to mark, it consults the table to find the link number assigned to the inbound interface. The core router uses the modulo technique for marking is calculated as in Equation 1,

New marking information= current marking information × number of interfaces on the router + the link number (1)

Reconstruct and Drop (RAD), which is built from the IDP and its location is one hop back from the victim. In general, the host follows the same path (shortest path) across the routers for sending the packet to destination. The RAD component maintains the information about each host and its equivalent packet marking value. When the marking value matches the stored value, it forwards the packet to respective host. At the time of the attack, when host spoofs the IP address of another host, the packet marking value differs from the value stored in the RAD. This is because, for marking CLASSIE uses MAC address instead of the IP address. Hence the packets are dropped at the victim side and RAD requests for the traceback.

When the victim is under the attack, it issues traceback request containing the marking information of the packet to be traced to the nearest router that delivers the packet. The upstream router uses the reverse modulo to find the inbound interface of the traceback requested packet using the marking information found in the traceback request and then using the hardware address table at the inbound interface, the router finds the previous upstream router connected to that interface.

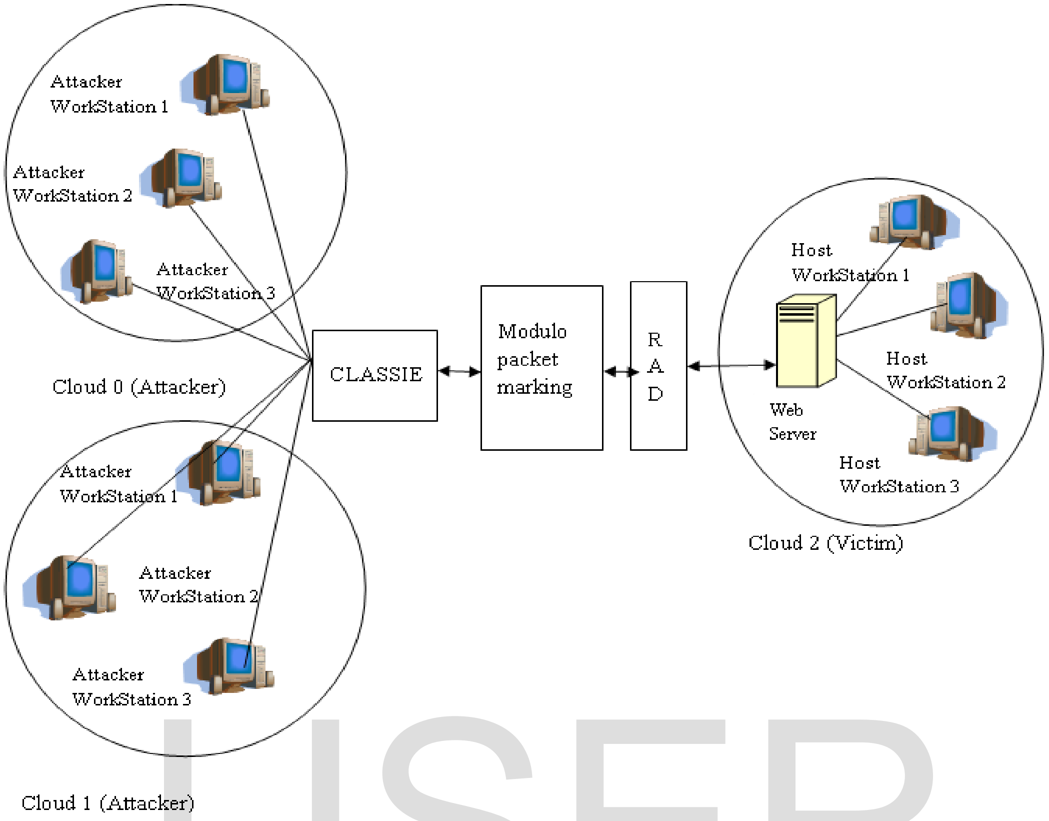

Fig.1. Conceptual diagram of proposed approach

Then, the upstream router becomes the current router and traceback procedure is repeatedly performed till the edge router of the sending host is reached. When this is done, the victim would have found the routers crossed by the attack packet and would send a request to the edge router to find the physical address of the node that originated the attack packet. RAD works by observes incoming messages and makes a decision about either allowing the message through or dropping it. It avoids the spoofing attack, as it finds the true origin of a packet.

4 Results and Discussion

The two important parameters used to measure the detection and filtering of DDoS attacks are Detection Rate and False Alarm Rate. Detection Rate (DR) of the attack traffic that is trained and tested by CLASSIE is equal to true positive (TP). True Positive is defined when a system gets an alert when an attack has taken place. False Positiveis defined when system gets an alert when no attacks have taken place. True Negative is defined as there is no attack from intruders as well as no alert from alarm. And the False Negative is defined as attacks have taken place in the system, but failed to detect them.

The comparison of the average length of the code required by the proposed model with the length required by the Huffman code is shown in Fig. 2 It shows that modulo packet marking requires lesser bits than Huffman way of coding.

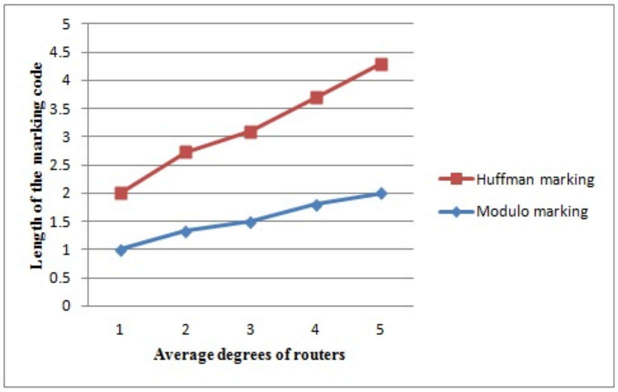

Fig. 2. Average length of the marking field for Huffman marking vs. Modulo marking

The proposed work was tested with different number of packets. The performance of the modulo packet marking increases, in detecting and trace backing the attack packets, when compared with the existing cloud protector as shown in Fig. 3.

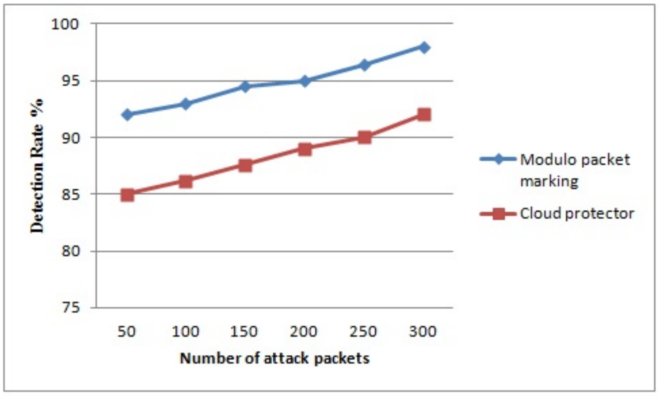

Fig. 3. Comparison of modulo packet marking with the cloud protector

5 Conclusion and Futurework

One of the most serious threats to cloud computing comes from HTTP or XML-Based DoS attacks. These attacks can be efficiently detected by using packet based marking approach on the attacker side and the detected packets are filtered by dropping the marked packets on the victim side. So, the packet marking overhead and the false positive rate of DoS attacks are greatly reduced. The detection of DDoS attack is improved by replacing the Cloud Protector with RAD on the victim side and the introduction of CLASSIE and modulo marking at the source side. This improves the reduction of the false positive rate and increase the detection and filtering of DDoS attacks. The future work can be extended by integrating the proposed system with the source end defensive systems to detect on MAC spoofing.

References

[1] A.Belenky and N.Ansari (2003), 'Tracing Multiple Attackers with Deterministic Packet Marking (DPM)', Proceedings of IEEE Pacific Rim conference on communications, computers and signal processing, Vol. 1, pp. 49–52.

[2] A.Chonka W. Zhou and Y.Xiang (2008a), 'Protecting Web Services with Service Oriented Traceback Architecture', Proceedings of the IEEE eighth international conference on computer and information technology, pp. 706-711.

[3] A.Chonka, W.Zhou and Y.Xiang (2008b), 'Protecting Web Services from DDoS Attacks by SOTA', Proceedings of the IEEE fifth international conference on information technology and applications, pp. 1-6.

[4] A.Chonka, W.Zhou, J.Singh and Y.Xiang (2008c), 'Detecting and Tracing DDoS Attacks by Intelligent Decision Prototype', Proceedings of the IEEE International Conference on Pervasive Computing and Communications, pp. 578-583.

[5] A.Chonka, W.Zhou and Y.Xiang (2009a), 'Defending Grid Web Services from X-DoS Attacks by SOTA', Proceedings of the third IEEE international workshop on web and pervasive security (WPS 2009), pp. 1-6.

[6] A.Chonka, W.Zhou and J.Singh (2009b), 'Chaos Theory Based Detection against Network Mimicking DDoS Attacks', Journals of IEEE Communications Letters, Vol. 13, No. 9, pp. 717-719.

[7] A.Chonka, Y.Xiang, W.Zhou and A.Bonti (2011), 'Cloud Security Defence to Protect Cloud Computing against HTTP-DoS and XML-DoS attacks', Journal of Network and Computer Applications, Vol. 34, No. 4, pp. 1097-1107.

[8] D.Dean (2002), 'An algebraic Approach to IP traceback', Journal ACM Transactions on Information and System Security', Vol. 5, No. 2, pp.119-137.

[9] S.Savage, D.Wetherall, A.Karlin and T.Anderson (2000), 'Practical Network Support for IP traceback', Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, pp. 295-306.

[10] H.Shabeeb, N.Jeyanthi and S.N.Iyengar (2012), 'A Study on Security Threats in Clouds', Journal of Cloud Computing and Services Science, Vol. 1, No. 3, pp. 84-88.

[11] X.Xiang, W.Zhou and M.Guo (2009), 'Flexible Deterministic Packet Marking: an IP Traceback System to Find The Real Source of Attacks', Journal of IEEE Transactions on Parallel and Distributed Systems, Vol. 20, No. 4, pp. 567-580.

[12] K.H.Choi and H.K.Dai (2004), 'A Marking Scheme using Huffman Codes for IP Traceback', Proceeding of 7th International Symposium on Parallel Architectures, Algorithms and Networks (SPAN'04).